Estimating models with network interactions and unobserved heterogeneity

Conference of the Italian Society of Statiscs (SIS), Palermo June 2018.

Luisa Corrado* and Salvatore Di Novo†

This Version: February 2018

Abstract

There is a growing awareness of the importance of connectivity in modelling which extends beyond its roots in spatial econometrics. Nodes need not necessarily be places and there is a burgeoning wider literature documenting the role of network interaction in affecting individual

*University of Rome Tor Vergata, Department of Economics and Finance, Italy. E-mail: luisa.corrado@uniroma2.it.

†University of Rome Tor Vergata, Department of Economics and Finance, Italy. E-mail: salvatore.di.novo@uniroma2.it.

decisions. In a variety of applications, the structure of networks is thought to influence final outcomes, and so it is important to understand how the networks themselves are formed and how to account for the latent factors affecting their formation in the final estimation of endogenous effects.

In this paper, we propose an approach to estimate models with network interactions in the presence of individual unobserved heterogeneity. The latter may impact the formation of ties and/or exogenous effects, thereby undermining identification of the associated parameters. In a panel setting, we devise a way to cope with these sources of endogeneity by relying on observable variations. When exogenous effects are involved, one can control for unobserved heterogeneity by including time-averages of the endogenous variables. When unobserved individual traits affect the process of network formation, it is possible to explore the role of network statistics affected by those unobserved traits. This is desirable and probably unavoidable, at least from a dimensionality standpoint. Indeed, to manage each link as standalone, one would need many data points as the number of potential links. This is pointless, as long as one has an easy and effective alternative at hand.

Such possibility seems to be not fully explored yet. Still, the network literature has emphasized the importance of network statistics, though it is actually crucial to focus on statistics displaying variation over individuals. We derive a 2SLS estimator where we address simultaneity bias using instruments provided by the product between successive powers of the network

matrix and the matrix of exogenous covariates and then complement the identification strategy with the control approach on the network statistics as highlighted above in order to cope with unobserved heterogeneity. The method we propose is also easy to implement as long as one has information on the structure of the relevant network. In this respect, it is worth stressing that, in principle, one does not need to have full access to entire structure of links, as availability of the aforementioned statistics would suffice as long as one has some (consistent) prior on the relevant network. In a panel setting, we assess the performances of our method via a Monte Carlo exercise, in which we consider various combination of models and intensities of network interactions. We further account for different intensities of the social multiplier and separately assess the cases in which unobserved sources hit the network structure only or act on exogenous effects as well. Focusing on the former case, our approach may be also applied when a simple cross-section is available, which makes it even more valuable under data constraints.

Key-Words: Networks, Individual Unobserved Heterogeneity, Dynamic Network.